



UNIVERSIDAD DE ZARAGOZA  
CENTRO POLITÉCNICO SUPERIOR



XVIII CONVOCATORIA DE PREMIOS  
"INGENIEROS DE TELECOMUNICACIÓN"

PREMIO BOSCH SECURITY SYSTEMS

# DESARROLLO DE UNA APLICACIÓN WEB DESTINADA A FACILITAR EL RECONOCIMIENTO FACIAL EN MALAS CONDICIONES DE ILUMINACIÓN USANDO UN ENTORNO VIRTUAL EN 3D

**AUTORA:** RAQUEL BLASCO LÁZARO

**DIRECTOR:** MARTIN TURNER

**PONENTE:** CARLOS ORRITE URUÑUELA

SEPTIEMBRE 2007

The University  
of Manchester

MANCHESTER  
1824



THE UNIVERSITY OF MANCHESTER  
MANCHESTER VISUALIZATION CENTRE

# Índice de Contenidos

1	Trabajo Realizado .....	1
1.1	Origen.....	1
1.2	Objetivos .....	2
1.3	Desarrollo .....	3
1.3.1	Modelado de las caras en 3D.....	3
1.3.2	Interfaz Gráfico de Usuario (GUI).....	3
1.3.3	Panel de Control e Iluminación .....	4
1.3.4	Medidas Numéricas.....	4
1.4	Conclusiones.....	6
2	Originalidad del Proyecto .....	7
3	Resultados.....	9
3.1	Análisis de las Caras Modeladas en 3D .....	9
3.2	Análisis de la Aplicación y su GUI.....	10
3.3	Análisis de la Comparación de Imágenes de Vídeo con la Cara en 3D .....	11
3.4	Evaluación Comparativa entre VRML y POV-Ray .....	11
4	Aplicabilidad Práctica .....	12
4.1	Aplicación Web.....	12
4.2	Modelos 3D en el <i>Visualization Immersive Programme Laboratory</i> .....	13
	Glosario de Acrónimos .....	14
	Bibliografía.....	15

# Índice de Figuras

Figura 2.1:	Comparación: Imagen de Vídeo - Fotografía .....	7
Figura 2.2:	Similaridad Facial en Función de la Diferencia en el Punto de Vista de los Sujetos .....	7
Figura 3.1:	Definición de las Medidas Faciales Utilizadas en el Análisis .....	9
Figura 3.2:	Porcentajes Medios de Similaridad .....	10
Figura 3.3:	Persona 1 – Porcentaje Similaridad .....	10
Figura 3.4:	Porcentaje de Error Relativo en la Herramienta de Medidas .....	10
Figura 3.5:	Variación de Ratios y su Similaridad en Función de la Orientación de la Cara .....	11
Figura 4.1:	Interfaz Gráfico de Usuario .....	12
Figura 4.2:	Vista del VIPL .....	13
Figura 4.3:	Análisis en el VIPL .....	13

# 1 Trabajo Realizado

## 1.1 Origen

La identificación facial en casos judiciales es una complicada tarea. Obtener imágenes del delincuente en las que se distinga su cara con claridad para poder utilizarlas como método de identificación suele ser complicado, ya que normalmente los criminales no miran directamente a la cámara de seguridad, sino que tratan de ocultar sus caras. Además, es frecuente que la iluminación de las imágenes capturadas no sea la más adecuada. La presencia, cada vez mayor, de cámaras de vigilancia en lugares que requieren altos niveles de seguridad, ha incrementado el uso de estas secuencias de vídeo como principal prueba en casos judiciales. En principio, esto podría reducir la frecuencia de errores humanos cometidos en identificación facial, ya que por ejemplo para un testigo, el cambio de contexto y una situación de alto estrés pueden influir en el proceso de identificación. Sin embargo aunque esto solucione algunos problemas, otros nuevos han aparecido y deben ser tratados.

Las imágenes grabadas y disponibles en procesos judiciales tienen normalmente una calidad de nivel medio-bajo y a veces extremadamente pobre. Varios aspectos reducen la utilidad de las imágenes capturadas y hacen que el proceso de identificación se complique. Para reducir costes de almacenamiento las imágenes suelen comprimirse, lo que significa pérdida de información y un posible degradado de la imagen más clara de la cara del delincuente. Por otro lado, el área cubierta por la cámara a veces es demasiado amplia, lo que supone imágenes de la cara del criminal con una resolución bastante baja, tomadas desde un ángulo no adecuado y afectadas por una iluminación no demasiado óptima. Los sistemas digitales están resolviendo algunos de estos problemas, pero no todos. Su aceptación como prueba está sujeta a la garantía de que el material es original <sup>[1], [2]</sup>.

En resumen, las cámaras de seguridad no resuelven por sí solas todos los problemas de identificación. La comparación de estas imágenes con fotografías de sospechosos puede llegar a ser una tarea complicada. Las imágenes suelen estar tomadas desde ángulos diferentes, lo que complica la extracción directa de medidas comparativas. Otros factores adicionales, como la variación de la luz y de las expresiones, incrementan el problema <sup>[3]</sup>. Por ello, es necesario que las secuencias de vídeo sean analizadas por analistas de imágenes expertos para reducir el riesgo de una interpretación incorrecta. Para facilitar este análisis se han desarrollado numerosas técnicas basadas en comparación de imágenes faciales. Sin embargo, todavía no ha sido estandarizado de forma universal ningún procedimiento.

Este tipo de comparaciones es posiblemente uno de los más difíciles tipos de identificación. Uno de los problemas más restrictivos de los sistemas de identificación facial en 2-dimensiones (2D) es la diferencia de pose, además de las sombras o de los diferentes tipos de expresión. Idealmente, para llevar a cabo la identificación independientemente de la pose, se debería disponer de imágenes de ambos sujetos desde todos los posibles ángulos de vista. Sin embargo, esto no es viable en una investigación policial. Este proyecto propone solucionar este problema utilizando un sistema de medida sobre caras en 3-dimensiones (3D).

El uso de modelos faciales en 3D permite obtener poses similares a las del sujeto del vídeo e incluso aplicar una iluminación similar para crear las mismas condiciones y facilitar el trabajo del analista. Las dimensiones faciales y las sombras serían entonces comparables, facilitando la aplicación de las técnicas de análisis <sup>[3]</sup>. Además un modelo facial en 3D aporta información de relieve, difícil de discernir en imágenes en 2D, especialmente frontales.

La *Unit of Art in Medicine* <sup>[4]</sup>, perteneciente a *The University of Manchester* <sup>[5]</sup>, ha trabajado e investigado nuevas técnicas de identificación facial en colaboración con la policía durante diecisiete años <sup>[1], [6], [7]</sup>. En esta ocasión, propuso al *Manchester Visualization Centre* (MCV) <sup>[8]</sup>, iniciar una investigación de cómo el uso de modelos 3D podría facilitar el trabajo de los analistas. Su propuesta fue crear un sistema de comparación interactivo que permitiese visualizar la cara en 3D del sospechoso y variar su pose e iluminación para poder compararla con la imagen de vídeo de un sujeto (el delincuente) mediante diferentes técnicas de comparación de imágenes.

## 1.2 Objetivos

El objetivo de este proyecto ha sido crear y evaluar una aplicación Web que facilite el trabajo de los analistas de imágenes que colaboran con la policía cuando un sospechoso es detenido en base a las imágenes de una cámara de seguridad. El método consiste en crear y modelar una cara en 3D para ajustarse de forma precisa a una imagen conocida del sospechoso. Este modelo en 3D puede ser entonces exportado a la aplicación creada, donde es posible aplicar una serie de transformaciones para crear una pose y unas condiciones de iluminación similares a las de una imagen del criminal tomada de una secuencia de vídeo. De esta manera, cuando la escala, orientación e iluminación de ambas imágenes sean las mismas o muy similares, será más fácil verificar si ambas pertenecen a la misma persona, ya que las dimensiones y formas faciales, así como las sombras generadas en la cara serán comparables. Además la herramienta permite extraer distancias y ratios faciales tanto en las imágenes en 2D como en el modelo 3D para analizar su similitud.

De forma paralela, se ha evaluado el uso de los productos de animación: Poser 6.0 <sup>[9]</sup> desarrollado por e-frontier <sup>[10]</sup> y FaceShop Pro <sup>[11]</sup> creado por Pantomat <sup>[12]</sup>, para crear modelos faciales en 3D de personas reales y utilizarlos dentro de la herramienta creada para llevar a cabo procesos de identificación facial.

A continuación se describen los principales objetivos de este Proyecto Fin de Carrera (PFC):

- Implementar una aplicación Web, de carácter open-source. Pudiendo ser utilizada en cualquier sitio con conexión a Internet, sin necesidad de comprar ningún software comercial.
- Diseñar un Interfaz Gráfico de Usuario (GUI) basado en Java <sup>[13], [14]</sup> que incluya:
  - Entorno en 3D creado en *Virtual Reality Modelling Language* (VRML) <sup>[15], [16]</sup> que contenga la cara en 3D del sospechoso y permita su manejo en tiempo real.
  - Secuencia de vídeo del delito para su comparación directa con el modelo 3D.
  - Múltiples e intuitivos modelos de iluminación controlables por el usuario para poder ajustar la iluminación en el entorno en 3D con la existente en la secuencia de vídeo.
  - Funciones de rotación, translación y escalado para que el usuario pueda ajustar la pose de la cara en 3D con la pose del sujeto de la secuencia de vídeo en tiempo real.
  - Herramienta de medida que permita la obtención y comparación de distancias y ratios faciales en el espacio 2D y 3D (imágenes de vídeo, fotografías y modelos en 3D) para poder realizar una comparación interactiva.
- Evaluar distintas herramientas de animación en 3D para generar caras en 3D de personas reales, con idénticas características y dimensiones faciales.
- Realizar un análisis comparativo entre los posibles lenguajes de programación disponibles en el mercado para visualizar las caras en 3D e interactuar con ellas.
- Crear una pequeña base de datos de varias personas con sus respectivas fotografías, modelos faciales en 3D, así como varias secuencias de vídeo, en diferentes escenarios y con diferentes condiciones para el análisis de las herramientas de animación 3D y para la evaluación de la aplicación Web.

## 1.3 Desarrollo

En este apartado se describe el trabajo que se ha llevado a cabo en este PFC.

### 1.3.1 Modelado de las caras en 3D

El uso de modelos faciales en 3D pretende solucionar algunos de los problemas que pueden darse en las comparaciones de imágenes faciales en 2D, como por ejemplo la diferencia en pose, escala o iluminación. Estas diferencias complican el análisis de las características faciales y la comparación de dimensiones y proporciones. Sin embargo, en un modelo en 3D, la pose y la escala pueden ser ajustadas. En este proyecto se han utilizado dos herramientas para modelar las caras en 3D.

- **Poser 6.0** permite crear una cara en 3D a partir de un modelo facial utilizado como plantilla. La textura facial es extraída a partir de dos fotografías, una frontal y otra lateral. La estructura de la cara, así como sus dimensiones, deben ser modificadas por el usuario usando una serie de controles que permiten modificar la morfología del modelo.
- **FaceShop Pro** genera una cara en 3D a partir de una sola fotografía. La aplicación requiere que el usuario indique ciertos puntos clave para calcular las dimensiones y proporciones faciales.

Ambas herramientas han sido evaluadas en función del tiempo requerido para completar el proceso, la precisión obtenida, la facilidad de uso y la flexibilidad de la herramienta, y la compatibilidad del formato de la salida.

FaceShop Pro, a diferencia de Poser, ajusta la textura y genera la estructura facial en 3D automáticamente. Es capaz de usar la información de la fotografía para calcular las proporciones faciales. Sin embargo, el resultado es estático y sus características no pueden ser retocadas manualmente. Poser ofrece más posibilidades, crea un modelo flexible y dinámico permitiendo incluso modificar la expresión facial del rostro. Además, es capaz de exportar los ficheros en una amplia variedad de formatos, entre ellos VRML. Sin embargo, el proceso de modelado es demasiado largo y complicado.

Por estos motivos, la solución tomada ha sido utilizar ambas herramientas en conjunto. FaceShop Pro es capaz de crear un buen modelo de partida que puede ser mejorado con Poser. De esta manera, el modelo puede llegar a ser más preciso, y el tiempo necesario es reducido.

### 1.3.2 Interfaz Gráfico de Usuario (GUI)

La aplicación requiere un interfaz gráfico que permita disponer de toda la información en un sólo entorno para poder analizarla y compararla. Este GUI debe ser intuitivo y fácil de manejar. Como se ha explicado previamente, la información consiste en fotografías, vídeos y modelos faciales en 3D. Los modelos 3D son mostrados dentro de un entorno virtual VRML en la ventana de *Cortona VRML Client*<sup>[17]</sup> y el GUI ha sido creado utilizando el API Java debido a su flexibilidad y a la multitud de posibilidades que ofrecen sus componentes.

VRML ha sido elegido frente a otros lenguajes de programación como POV-Ray<sup>[18]</sup>, debido a que permite interactuar con las escenas en tiempo real.

La comunicación entre los *applets* de Java y VRML no es directa. Es necesario utilizar funciones JavaScript embebidas dentro del fichero HTML, que serán llamadas cuando se realicen cambios en los *applets* o se produzca algún evento en el entorno 3D. Estas funciones son capaces de modificar las variables de los nodos contenidos en el fichero VRML, así como de transmitir información al GUI de Java.

### 1.3.3 Panel de Control e Iluminación

Para poder ajustar la pose del modelo 3D con la pose del sujeto de la secuencia de vídeo es necesario incluir funciones de rotación, translación y escalado en la aplicación. En la aplicación Web se ha incluido un Panel de Control e Iluminación para interactuar con el modelo 3D del sospechoso.

VRML representa la rotación en base al método *Axis Angle*. Este método utiliza cuatro números. Los tres primeros definen una posición en el espacio de coordenadas 3D, definiendo el eje de rotación como la línea que une el origen y esa posición. El cuarto número indica el número de radianes que debe rotarse el objeto alrededor de ese eje <sup>[16]</sup>.

Dado que este método de rotación no es muy intuitivo para el usuario, se han incluido en el GUI tres deslizadores que permiten rotar el objeto según los ángulos de Euler ( $\alpha, \beta, \gamma$ ). La aplicación se encarga internamente de transformar estas rotaciones en términos legibles para VRML.

Además de incluir controles para manipular la pose del modelo en 3D, el GUI posee una serie de paneles que permiten variar las condiciones de iluminación en el entorno virtual VRML. Se han incluido varias fuentes de luz basándose en los modelos de iluminación de VRML: *point light*, *directional light*, *spotlight* y la iluminación por defecto *headlight*. Un panel de iluminación principal controla las fuentes activas en el entorno 3D y una serie de sub-paneles individuales correspondientes a cada fuente de luz permiten personalizar cada una de sus características individuales. Realizando combinaciones de estos modelos de iluminación es posible obtener diferentes efectos y sombras en el modelo 3D.

### 1.3.4 Medidas Numéricas

La aplicación incluye una Herramienta de Medidas que permite:

- Extraer medidas faciales comparativas en el modelo 3D y las fotografías del sospechoso para poder evaluar la precisión del modelo. Este método incluye un proceso de calibración con medidas reales para poder obtener las medidas en centímetros.
- Obtener un porcentaje de similitud entre los ratios faciales del modelo 3D del sospechoso y el sujeto de la secuencia de vídeo para determinar si podrían ser la misma persona, creando así una comparación independiente de la diferencia de escala y sin unidades métricas.

El primer sistema de medida consiste en extraer distancias faciales tanto en el modelo 3D como en las fotografías del sospechoso, seleccionando con el ratón las medidas que se quieren realizar.

La obtención de medidas comparativas requiere que sean tomadas con un mismo sistema de medida. Es necesario por tanto, fijar un sistema de referencia común entre las fotografías y el modelo 3D de VRML. Una herramienta ha sido incluida en el GUI para realizar este proceso de calibración. Esta herramienta es capaz de calcular la correspondencia entre las medidas extraídas del modelo 3D y de las fotografías con las medidas reales del sospechoso en centímetros. Para ello, el usuario debe indicar una serie de puntos faciales en el modelo 3D y en las fotografías, y a su vez, indicar en el Panel de Calibración cuales son las medidas faciales reales del sospechoso en centímetros. El sistema calculará y almacenará cada uno de los factores de conversión para poder devolver las sucesivas medidas en centímetros. Los puntos y distancias faciales que se han seleccionado para la calibración han sido elegidos en base a su estabilidad frente a cambios de expresión.

Una vez realizado este proceso de calibración, la Herramienta de Medidas habilita dos paneles que permiten extraer distancias numéricas en centímetros, tanto del modelo 3D, como de las fotografías del sospechoso. En ambos casos es el usuario quien debe señalar los puntos de medida para obtener sus coordenadas y calcular su distancia Euclídea.

En el modelo 3D la distancia Euclídea entre los puntos seleccionados es calculada tanto en el sistema de coordenadas 3D, como en el plano XY (2D), paralelo a la pantalla del ordenador. De esta forma, los resultados en 3D pueden ser comparados con las medidas tomadas del sospechoso en persona, y los resultados en 2D pueden ser utilizados para realizar comparaciones con las medidas tomadas en las fotografías. Para realizar el cálculo de la proyección lineal en 2D es necesario conocer la orientación del modelo 3D, ya que las coordenadas obtenidas en VRML son independientes de la pose. El Panel de Control e Iluminación envía los ángulos de Euler cada vez que son modificados, de forma que estén siempre actualizados. A partir de ellos se obtiene la matriz de rotación para poder calcular las coordenadas rotadas de cada uno de los puntos. La distancia en 2D es entonces obtenida calculando la distancia Euclídea entre la proyección ortogonal de ambos puntos.

El segundo sistema de medida que se ha incluido en la Herramienta de Medidas consiste en utilizar ratios faciales como sistema de comparación. Se ha incluido un panel que permite calcular el porcentaje de similaridad entre los ratios del modelo 3D con la imagen del sujeto de la secuencia del delito para determinar la probabilidad de que sean la misma persona. Hay que tener en cuenta que un porcentaje muy elevado no implica necesariamente que sean el mismo sujeto, ya que existen individuos que poseen ratios faciales similares. Sin embargo, un porcentaje de similaridad pequeño podría descartar al sospechoso.

Para diseñarlo se han evaluado dos consideraciones:

- Ambos sujetos podrían ser o no la misma persona. El objetivo es comparar sus dimensiones faciales. El proceso de calibración no puede ser aplicado a la imagen de vídeo porque las distancias faciales del sujeto son desconocidas. Sin embargo, la comparación puede realizarse usando ratios faciales, ya que su medida no requiere calibración.
- Las imágenes capturadas por una *Closed Circuit Television* (CCTV) son a menudo no frontales. Utilizando el modelo 3D puede deducirse la matriz de rotación que afecta a la imagen de vídeo, para poder aplicarle su inversa y obtener así los ratios correspondientes al plano frontal.

Se trata de un proceso guiado que indica en todo momento los pasos que debe seguir el usuario. El GUI incluye una imagen de un modelo facial donde se indican los puntos que deben ser señalados. Por defecto se indica como medir el ratio entre la distancia exterior de los ojos y la longitud de la nariz, ya que la variación con las expresiones en estas medidas es relativamente pequeña. Sin embargo, la herramienta permite calcular el ratio entre cualquier par de distancias si es necesario.

Estas distancias serán calculadas en centímetros si la calibración ha sido realizada, o en unidades genéricas si no ha sido así. En ambos casos, su ratio será calculado y enviado junto con ambas distancias para ser mostrados en el interfaz.

Las condiciones de escalado y pose en una imagen en 2D son fijas y en principio desconocidas. Sin embargo, estas condiciones se conocen en el modelo 3D. Su orientación y escalado pueden ser ajustadas para obtener condiciones similares a las del sujeto que aparece en el vídeo. De esta forma, puede deducirse la matriz de rotación a partir de los ángulos de Euler que afectan al modelo 3D, ya que la rotación sufrida por éste es la misma que la sufrida por el sujeto de la imagen del vídeo. Una vez obtenida, los puntos faciales marcados en la imagen del vídeo pueden ser trasladados al plano frontal, multiplicando sus coordenadas por la matriz de rotación inversa. Las distancias Euclídeas entre los puntos pueden ser entonces calculadas y utilizadas para obtener el ratio facial.

Los puntos marcados en la cara en 3D no necesitan ser rotados, ya que las coordenadas obtenidas en el entorno 3D de VRML son independientes de la orientación. Sin embargo, deben ser proyectados al plano XY antes de que la distancia Euclídea sea calculada. Cuando ambos ratios están disponibles, el porcentaje de similaridad es calculado y mostrado en el interfaz.

La herramienta no permite hacer comparaciones de distancias y de ratios al mismo tiempo. Sin embargo, sí que es posible extraer simultáneamente medidas en el modelo 3D y en las imágenes del vídeo para poder comparar resultados.



## 1.4 Conclusiones

Se ha creado una aplicación Web para facilitar el trabajo de los analistas de imágenes en el proceso de identificación facial de sospechosos a partir de imágenes de vídeo procedentes de cámaras CCTV. Esta aplicación engloba en un único interfaz toda la información disponible, permitiendo interactuar con ella y compararla.

Esta aplicación ofrece la posibilidad de trabajar con un modelo facial 3D del sospechoso. La herramienta permite manipular la pose y el escalado de la cara en 3D, además de incluir modelos de iluminación dinámicos. De esta forma pueden ajustarse para obtener condiciones similares a las del sujeto de la imagen de vídeo, y facilitar así la comparación. Esto soluciona los problemas de diferencia de pose, escalado e iluminación que afectan a la comparación de imágenes en 2D.

Además la aplicación incluye una Herramienta de Medidas que permite extraer y comparar distancias y ratios faciales, en las fotografías y el modelo 3D del sospechoso, y en las imágenes capturadas por una cámara CCTV. La herramienta puede ser calibrada para mostrar todas las medidas en centímetros, y facilitar su comparación. Para determinar el porcentaje de similaridad entre las características faciales del sospechoso y de la persona que aparece en el vídeo, pueden realizarse comparaciones de ratios faciales. Éstos son independientes de la diferencia de escalado y no requieren calibrado ya que no tienen unidades métricas. Una diferencia suficientemente clara entre ratios sería suficiente para descartar al sospechoso.

Para evaluar la eficiencia de la aplicación se ha creado una pequeña base de datos de ocho personas. Esta contiene una serie de fotografías, vídeos, distancias faciales y caras en 3D.

El modelado de las caras en 3D ha sido realizado con Poser y FaceShop Pro, ya que uno de los objetivos era evaluar su eficiencia para crear modelos 3D de personas reales. Si ambas son usadas de forma conjunta, pueden obtenerse mejores resultados y el tiempo de creación se reduce. La precisión de todos los modelos ha sido evaluada positivamente con un porcentaje de similaridad medio del 96,394 %. Sin embargo, existen algunas limitaciones. El proceso de modelado es bastante laborioso y depende en cierto grado de las habilidades artísticas del usuario. La precisión de las dimensiones y de la estructura facial suele ser mejor en la vista frontal que en el resto. Además, y pese a la flexibilidad de Poser, el modelado de la estructura facial está limitado a una serie de estructuras, y no es posible crear determinados rasgos faciales.

El funcionamiento de la aplicación ha sido evaluado positivamente. La comunicación entre los diferentes paneles de su interfaz gráfico es correcta y se realiza en tiempo real. Las medidas realizadas con la Herramienta de Medidas tienen un porcentaje de exactitud de al menos un 90 % y se ha evaluado positivamente su precisión al extraer y comparar ratios faciales entre diferentes imágenes de vídeo y el modelo 3D. Sin embargo, esta precisión depende en cierto grado de la pose del sujeto, ya que los puntos de medida deben ser claramente visibles. Finalmente, se ha comprobado que el Panel de Control e Iluminación permite ajustar la pose y la iluminación en el modelo 3D de la forma deseada.

Se ha realizado un estudio comparativo entre VRML y POV-Ray. POV-Ray crea imágenes de alta resolución con efectos realistas de iluminación, pero VRML no es capaz de crear estos efectos. Esto es debido a que el principal objetivo de VRML es poder interactuar con las escenas y para ello el tiempo de procesado de las imágenes debe ser pequeño. En POV-Ray, este tiempo es demasiado elevado. Por ello, se ha utilizado VRML en el desarrollo de la aplicación y el uso de POV-Ray se ha limitado a poder visualizar imágenes estáticas del modelo 3D de alta resolución.

Este proyecto va a ser utilizado en un entorno real. Su uso facilitará el trabajo de comparación de imágenes a los analistas de la *Unit of Art in Medicine* en identificación facial de sospechosos detenidos en base a las imágenes tomadas por una cámara de seguridad.

## 2 Originalidad del Proyecto

La comparación de imágenes procedentes de cámaras de seguridad con fotografías de sospechosos puede llegar a ser una tarea complicada (Figura 2.1). Las imágenes disponibles suelen tener una calidad bastante pobre, con baja resolución y mala iluminación. Sin embargo el reciente incremento del uso de cámaras de vigilancia para asegurar la seguridad ciudadana, ha incrementado también el número de juicios donde estos vídeos se presentan como prueba incriminatoria y es necesario investigar nuevos métodos para facilitar el trabajo de los analistas de imágenes.



Figura 2.1: Comparación: Imagen de Vídeo - Fotografía

Uno de los problemas más restrictivos de los sistemas de identificación facial en 2D es la diferencia de pose. Las imágenes suelen estar tomadas desde diferentes ángulos, lo que complica la extracción directa de medidas comparativas. Otros factores adicionales, como la variación de la luz y de las expresiones, incrementan el problema. La probabilidad de obtener una imagen nítida y frontal del sujeto que aparece en un vídeo de una cámara de seguridad es extremadamente pequeña.

El gráfico mostrado en la Figura 2.2 describe el grado de similitud entre imágenes faciales en función de la diferencia de ángulo visual. Cuanto mayor es la diferencia entre los puntos de vista, mayores son los errores cometidos. En determinados casos, incluso diferentes individuos con la misma pose podrían resultar más similares que la misma persona con una pose completamente diferente. Por este motivo, en identificación forense es necesario considerar este problema, ya que una baja similitud puede estar simplemente debida a una diferencia de pose <sup>[19]</sup>.

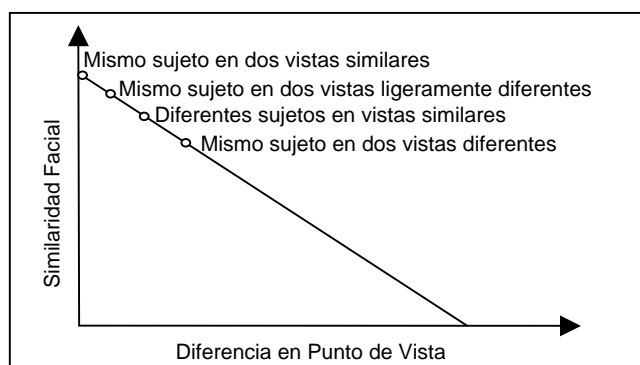


Figura 2.2: Similaridad Facial en Función de la Diferencia en el Punto de Vista de los Sujetos

Lo ideal sería disponer de imágenes de ambos sujetos desde todos los ángulos posibles, pero esto no es viable en una investigación policial. La originalidad de este proyecto radica en la propuesta de crear un sistema de comparación utilizando modelos en 3D. De esta forma se tiene información desde todos los posibles ángulos y es posible conseguir una pose similar a la del sujeto del vídeo, facilitando la aplicación de técnicas de análisis de imágenes, como por ejemplo, la superposición de imágenes, la alineación de características faciales o la extracción de medidas comparativas. Estos modelos además, aportan información de relieve permitiendo hacer análisis más exhaustivos que en imágenes en 2D, donde es algo difícil de discernir, especialmente en imágenes frontales.

Otro de los puntos novedosos que introduce este PFC es el uso de software de animación, tales como Poser 6.0 y FaceShop Pro para crear modelos 3D de personas reales y evaluar así, su posible utilidad en facilitar procesos de identificación criminal. El objetivo es utilizar estas herramientas para crear los modelos faciales en 3D de los sospechosos y exportarlos a la aplicación Web diseñada donde es posible variar su pose e iluminación y extraer medidas comparativas.

Existen otras técnicas para reconstruir modelos faciales en 3D, de diferente coste y con resultados de mayor o menor resolución. Por ejemplo un escáner láser permite obtener modelos de alta resolución, pero el coste del equipamiento suele ser bastante elevado. Otros métodos suponen un menor coste y una mayor portabilidad de sus sensores, como por ejemplo, utilizar dos o más imágenes estereoscópicas <sup>[20]</sup> o incluso usar imágenes en movimiento. Este caso en particular es interesante, ya que sólo requiere una cámara, no necesariamente calibrada, pero su principal problema es su baja resolución en cuanto al relieve <sup>[21]</sup>.

Todas estas técnicas crean modelos 3D estáticos. La mayor ventaja y el motivo principal por el que se ha propuesto este método ha sido que permite crear modelos dinámicos, es decir, una vez creado el modelo, es posible cambiar la expresión facial de la cara, lo que supone una innovación en las técnicas estudiadas hasta el momento. Esto supondría que no sólo podría obtenerse una iluminación y una pose similar a la del sujeto del vídeo, sino que además podría modelarse la misma expresión facial, haciendo comparativos a simple vista los rasgos faciales y facilitando la extracción de medidas comparativas directas. Lo que supone una gran ventaja frente al hecho de comparar distintas imágenes de un sujeto con un modelo 3D de expresión fija y neutral.

El objetivo ha sido crear una aplicación Web, accesible desde cualquier punto con conexión a Internet, con un GUI interactivo y de fácil manejo para el usuario, que recoja toda la información necesaria en un mismo entorno de forma que simplifique su comparación. Por ello, la opción elegida para crear el entorno en 3D ha sido VRML, ya que permite crear entornos 3D dinámicos y ofrece la posibilidad de poder interactuar con las escenas en tiempo real y poderlas visualizar desde cualquier punto de vista, lo que puede ser muy útil para el analista. Su gran ventaja es que posee modelos de iluminación que pueden ser modificados en tiempo real. Otros formatos como POV-Ray ofrecen una mayor calidad de imagen, pero debido a sus complejos procesos de cálculo matemático para calcular reflexiones y proyecciones de sombras no ofrecen interactividad con el usuario. Simplemente permiten fijar las condiciones en el modelo 3D y obtener imágenes 2D estáticas. Por ello, en este proyecto se propuso utilizar estos formatos sólo para obtener imágenes 2D de alta resolución en caso de que los analistas lo requiriesen.

La posibilidad de obtener una iluminación similar facilita el trabajo del analista, ya que no sólo las dimensiones son comparables, sino también las sombras pueden ser evaluadas debido a que el modelo ofrece información de relieve en 3D.

Otra de las funcionalidades que ofrece esta herramienta es que además de comparar las características faciales y los efectos de iluminación de un modo puramente visual, permite extraer y comparar medidas faciales tanto en el modelo 3D como en las fotografías 2D del sospechoso. De esta forma se puede analizar su similaridad y evaluar el grado de precisión del modelo 3D creado. Por otro lado, la herramienta ofrece la posibilidad de obtener un porcentaje de similaridad entre el modelo 3D del sospechoso y el sujeto que aparece en la secuencia de vídeo, a partir de la comparación sus ratios faciales. Este sistema de medida ha sido seleccionado entre otros ya que su metodología permite hacer una comparación de forma independiente a la diferencia de escalado y al sistema de medida.

Todas estas posibilidades que ofrece la aplicación tienen el objetivo de ayudar al analista a evaluar el grado de probabilidad de que el sospechoso sea la persona que aparece en el vídeo del delito. Su objetivo no es encontrar un elevado número de similitudes entre ambos sujetos, sino que encontrar diferencias significativas. Una clara diferencia es suficiente para probar que dos sujetos no son la misma persona. Sin embargo, varias similitudes no prueban que sean la misma. Por ejemplo, una clara diferencia en proporciones faciales, pecas, o lunares podría ser suficiente.

### 3 Resultados

Como principal resultado del proyecto cabe destacar la aplicación Web diseñada, dado que su uso pretende facilitar el trabajo de los analistas de imágenes. Sin embargo, esta sección está centrada en los resultados obtenidos al evaluar el correcto funcionamiento de la aplicación y la precisión de los modelos 3D creados.

#### 3.1 Análisis de las Caras Modeladas en 3D

La base de datos creada está compuesta por ocho personas. La información recogida consiste en varias fotografías, vídeos y una serie de medidas faciales. Esta información se ha utilizado para crear los modelos faciales en 3D de cada persona. La precisión de estos modelos ha sido evaluada mediante diferentes técnicas de análisis de imagen <sup>[1], [22]</sup>:

- **Alineación de Características Faciales:** Este método compara las dimensiones faciales alineando ambas imágenes en el plano horizontal y vertical, tomando como referencia determinadas características faciales.
- **Superposición de Imágenes:** La finalidad de esta técnica es evaluar la similaridad de las formas y proporciones faciales mediante sucesivas superposiciones de imágenes.
- **Comparación Proporcional:** Este método compara las distancias y los ratios faciales en ambas imágenes de forma numérica. La Herramienta de Medidas ha sido utilizada para extraer las distancias faciales (Figura 3.1).
- **Análisis en el Visualization Immersive Projection Laboratory (VIPL) <sup>[23]</sup>:** La cara en 3D ha sido transferida al laboratorio VIPL situado en el MVC. Este laboratorio está dotado de la tecnología necesaria para visualizar los datos en estéreo, de forma que el ojo humano perciba la imagen en 3D a tamaño real.

La evaluación de los resultados de todas estas pruebas a las que han sido sometidos los modelos 3D ha sido positiva. El gráfico de la Figura 3.2 muestra que el porcentaje de similaridad entre los modelos y las fotografías es de al menos 93,211%. La Figura 3.3 contiene un gráfico más detallado de uno de los sujetos, con el porcentaje de similaridad de cada medida. Este porcentaje es siempre mayor que 93,24% y su valor medio es de 97,38 %.

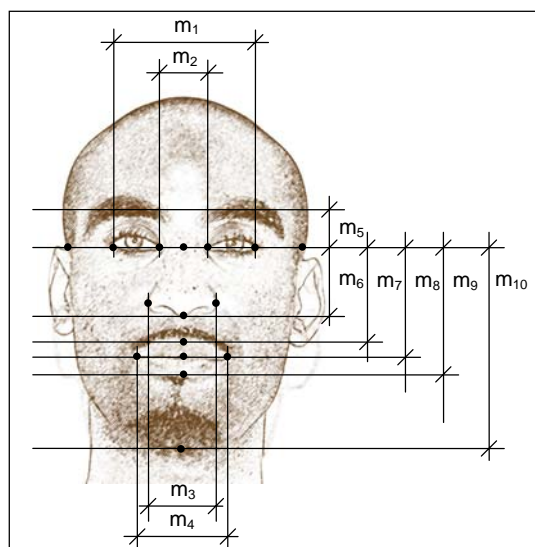


Figura 3.1: Definición de las Medidas Faciales Utilizadas en el Análisis

Sin embargo, a pesar de que los resultados obtenidos han sido evaluados positivamente, existen algunas limitaciones en el proceso de modelado de las caras en 3D. Se trata de un método bastante laborioso y que depende en cierto grado de las capacidades artísticas del usuario. Además, la precisión de las medidas es mejor en el plano frontal debido a que se

dispone de mayor información y por último, a pesar de la flexibilidad de los modelos de Poser, el modelado de las estructuras faciales está limitado a una serie de estructuras y no es posible crear determinados rasgos faciales.

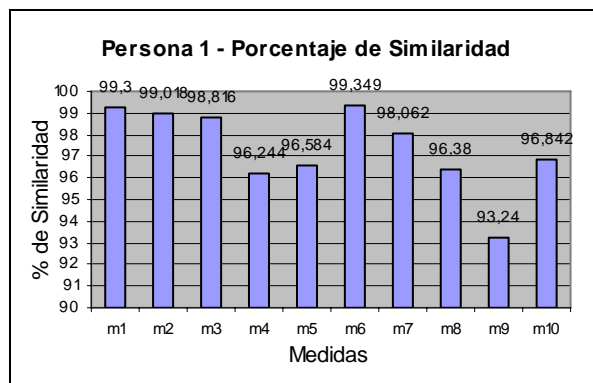
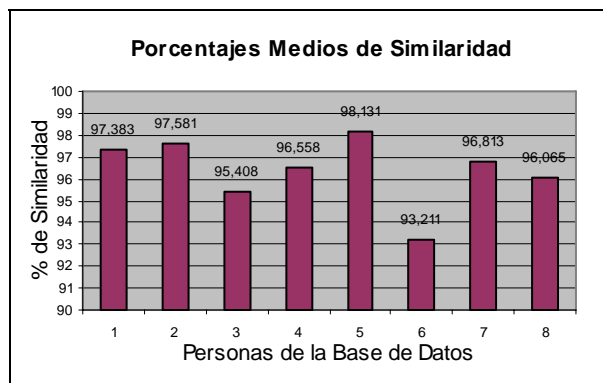


Figura 3.2: Porcentajes Medios de Similitud      Figura 3.3: Persona 1 – Porcentaje Similitud

### 3.2 Análisis de la Aplicación y su GUI

La aplicación y los componentes de su GUI han sido evaluados. Se han realizado diferentes pruebas para asegurar el correcto funcionamiento y la comunicación entre los diferentes componentes del interfaz en situaciones normales y en casos de excepción.

Por otro lado se ha evaluado la precisión de los resultados numéricos obtenidos en la Herramienta de Medidas. Este análisis ha consistido en calibrar la herramienta y extraer una serie de medidas (Figura 3.1) en las fotografías y en el modelo 3D. Cada una de estas distancias ha sido tomada diez veces para obtener medidas más fiables y disminuir el posible error humano en la selección de los puntos. Estas medidas han sido comparadas posteriormente con las distancias faciales del sujeto. El gráfico de la Figura 3.4 representa el porcentaje de error relativo en las medias tomadas en las fotografías y en el modelo 3D. Su valor medio puede observarse en la tabla de la Figura 3.4.

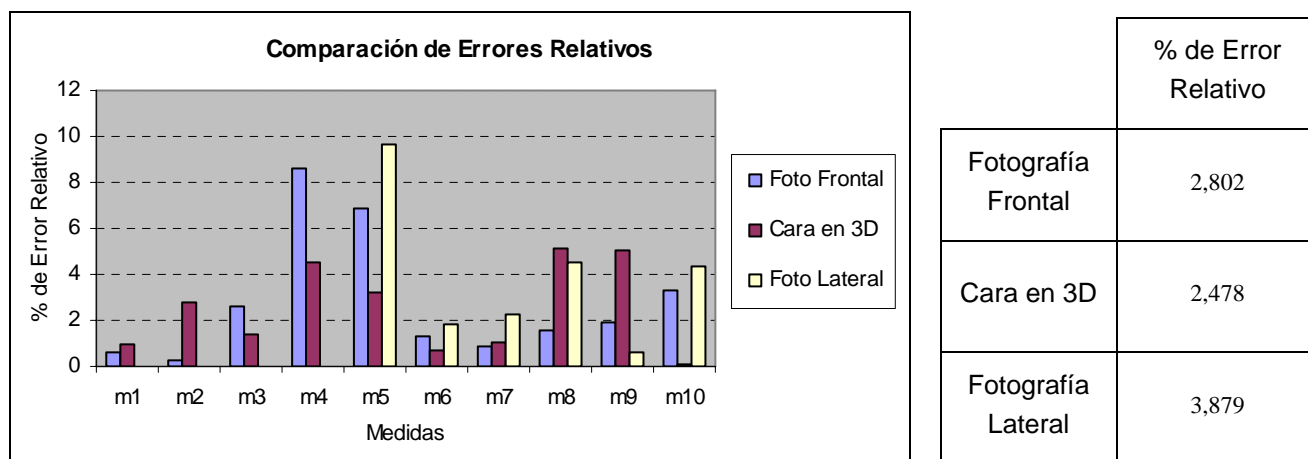


Figura 3.4: Porcentaje de Error Relativo en la Herramienta de Medidas

Todos los porcentajes de error están por debajo del 10 %. Esto significa que el porcentaje de precisión es al menos del 90%. Comparando los resultados se observa que las medidas en la fotografía lateral tienen un porcentaje de error mayor. Esto es debido a que el número de medidas para calibrarla es menor.

Las medidas utilizadas para la calibración de la herramienta han sido elegidas en función a su estabilidad frente a cambios de expresión y en función de cómo están definidos sus puntos de medida para hacer la selección. Los errores

cometidos en las medidas  $m_1$ ,  $m_2$  y  $m_6$  son muy pequeños. Están siempre por debajo de 1,32% en la fotografía frontal, 1,82% en la fotografía lateral y 2,769% en el modelo 3D. Sin embargo las medidas  $m_4$  y  $m_5$  tienen los mayores porcentajes de error. La distancia  $m_4$  corresponde a la longitud de la boca, la cual es especialmente sensible a los cambios de expresión y la medida  $m_5$  se corresponde con la distancia entre las cejas y los ojos. Ésta podría variar con la expresión, pero el error cometido podría estar también debido a que los puntos de medida no están muy bien definidos.

### 3.3 Análisis de la Comparación de Imágenes de Vídeo con la Cara en 3D

En este apartado se analiza la precisión de la Herramienta de Medidas para extraer y comparar ratios faciales entre las imágenes de vídeo y el modelo 3D.

El análisis ha consistido en extraer los ratios de un sujeto en diferentes imágenes de vídeo, y compararlos con los del modelo 3D. Con este método se pretende evaluar la influencia de la pose en la precisión de las medidas. Los gráficos de la Figura 3.5 muestran los ratios obtenidos en las imágenes de vídeo y en la cara en 3D, así como el porcentaje de similitud entre ellos. El valor medio del porcentaje de similitud entre ratios es del 86,511% y sus valores oscilan entre el 75% y el 99%. Los resultados han demostrado la influencia de la pose de la cara en la precisión de las medidas. La precisión depende del eje de rotación y de si el valor del ángulo de rotación es demasiado elevado. Rotaciones en las cuales los puntos de medida permanecen visibles no son problemáticas. Sin embargo, si la visibilidad de estos puntos es afectada, la precisión de las medidas se reduce.

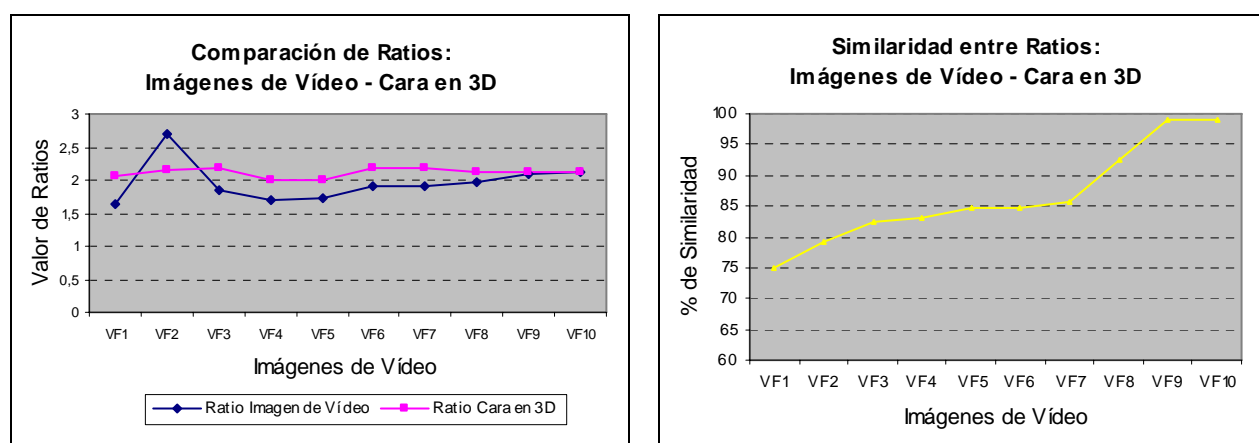


Figura 3.5: Variación de Ratios y su Similitud en Función de la Orientación de la Cara

La capacidad del Panel de Control e Iluminación para ajustar la pose y la iluminación en la cara en 3D ha sido también evaluada positivamente, ya que la pose del modelo 3D debe ser correctamente ajustada a la del sujeto de la imagen de vídeo para poder medir sus ratios faciales.

### 3.4 Evaluación Comparativa entre VRML y POV-Ray

Finalmente, se ha realizado un estudio comparativo entre VRML y POV-Ray. POV-Ray genera imágenes de alta resolución, con efectos realistas de iluminación. VRML es incapaz de crear estos efectos. Sin embargo, el uso de simples modelos de iluminación permite la interacción con las escenas en tiempo real. Este dinamismo no es posible en POV-Ray, dado que su proceso de trazado de imágenes es demasiado complejo. Requiere un tiempo de procesamiento demasiado elevado como para poder interactuar con las escenas. Además, este tiempo aumenta con la cantidad de iluminación. Dado que el objetivo es que el usuario pueda interactuar en tiempo real con el modelo 3D, se ha utilizado VRML para generar la escena. El uso de POV-Ray en este proyecto se ha limitado a crear una escena con la pose y la iluminación del modelo ajustadas a las de la imagen de vídeo; pudiendo generar a partir de ella, imágenes de alta resolución que pudiese utilizar posteriormente el analista de imágenes, pero sin la posibilidad de interactuar con ellas.

## 4 Aplicabilidad Práctica

La mayor aplicabilidad que tiene este proyecto es la utilidad de la aplicación diseñada para facilitar el trabajo de analistas de imágenes. Sin embargo, este proyecto ha constituido la primera fase de un proceso de investigación que va a ser continuado en el MVC de *The University of Manchester*. En los siguientes apartados se describe la aplicabilidad práctica de cada una de las ramas de investigación que ha desarrollado este proyecto.

### 4.1 Aplicación Web

La aplicación desarrollada en este PFC ha sido implantada y está siendo utilizada por los analistas de imágenes que trabajan en el departamento *Unit of Art in Medicine*, perteneciente a *The University of Manchester*. Este departamento está involucrado con la investigación de nuevas técnicas de identificación facial para su uso en casos judiciales, colaborando estrechamente con la policía.

Igualmente la aplicación (Figura 4.1) ha sido almacenada en un servidor del MCV para que pueda estar disponible a través de Internet para toda persona o institución que pueda requerir su uso, sin necesidad de adquirir ningún software comercial ([www.rcs.manchester.ac.uk/research/shadowimage/GUIMarley.html](http://www.rcs.manchester.ac.uk/research/shadowimage/GUIMarley.html)).

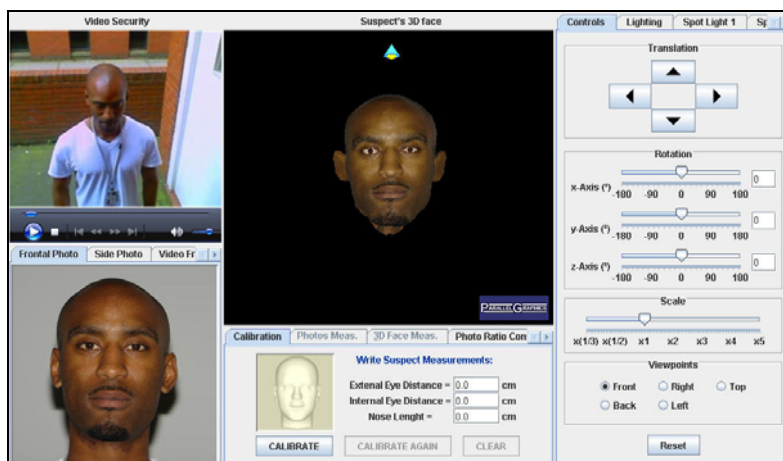


Figura 4.1: Interfaz Gráfico de Usuario

El objetivo es cambiar y facilitar el modo de trabajo de los analistas de imágenes. Hasta ahora su trabajo estaba limitado por numerosas restricciones que dificultaban el proceso y hacían de él una tarea lenta y complicada. Este sistema de comparación interactivo permite visualizar la cara en 3D del sospechoso y variar su pose e iluminación para poder compararla con la imagen de vídeo de un sujeto (el delincuente). De esta manera, se solucionan los problemas de pose derivados de las imágenes en 2D y los problemas referentes a la mala iluminación, pudiendo conseguir una pose y una iluminación idénticas en ambos sujetos. El analista puede comparar así fácilmente, las características faciales y los efectos de iluminación de un modo puramente visual, utilizando técnicas de alineación o de superposición de imágenes. Por otro lado, utilizando la Herramienta de Medidas es posible extraer y comparar medidas y ratios faciales, obteniendo un porcentaje de similitud, que permita evaluar la posibilidad de que ambos sujetos sean la misma persona. El uso de modelos en 3D mejoraría la eficiencia de los analistas, facilitando la aplicación de diferentes técnicas de análisis sobre las imágenes <sup>[1], [20]</sup>.

Por otro lado, el uso de estos modelos 3D y el dinamismo de Poser permiten personalizar la expresión del rostro de los modelos de forma que se facilita la comparación de los rasgos faciales y sus dimensiones.

Además la aplicación posee un interfaz de usuario interactivo y de fácil manejo para el usuario, de forma que el analista es capaz de utilizar la herramienta y manejar sus controles de manera intuitiva. Se han incluido en el GUI diferentes mensajes e imágenes de ayuda, además de incluir procesos guiados, mensajes de error y de poder controlar el interfaz tanto con el ratón como con el teclado.

Vista la aplicabilidad de este proyecto, el departamento *Unit of Art in Medicine* continúa trabajando en la mejora y ampliación de este proyecto. Actualmente se está llevando a cabo en el MVC un proyecto para ampliar las posibilidades de esta herramienta. Se trata del diseño de un proceso semi-automático para modelar caras en 3D a partir de datos de vídeo, de forma que pueda ser utilizado por usuarios no profesionales del mundo de las Tecnologías de la Información (IT). La tarea consistirá en crear la estructura facial a partir de un vídeo que posea información de la misma desde todos los puntos de vista. Este vídeo contará con indicadores de escala y posición. La salida será un fichero mejorado de VRML que se incluirá dentro de la herramienta actual para poder interactuar con él. Este proyecto pretende solucionar las actuales limitaciones del uso de herramientas de animación en el proceso de modelado de caras en 3D.

Finalmente, después de que la herramienta haya sido usada y evaluada durante un periodo de tiempo por los analistas, es posible que hagan sugerencias sobre como la herramienta podría ser mejorada. Podrían requerirse más opciones o más métodos de comparación, dependiendo de sus necesidades. Por ejemplo, podrían añadirse nuevos métodos de medidas como la comparación de proporciones, que consiste en seleccionar una distancia facial como el nivel de 100% y evaluar las posteriores medidas como proporciones respecto a esta referencia. También podrían automatizarse algunos de los procesos, haciendo que la herramienta fuese capaz de detectar determinados puntos faciales de forma automática y realizase los procesos de calibrado y de extracción del porcentaje de similitud de forma transparente para el usuario.

## 4.2 Modelos 3D en el *Visualization Immersive Programme Laboratory*

Uno de los métodos utilizados para evaluar imágenes es la estereoscopia. Consiste en mostrar sucesivas imágenes con diferentes ángulos de vista, de forma que el ojo humano perciba una imagen en 3D. Para ello se pueden utilizar unas gafas estereográficas que muestran a cada ojo dos imágenes con un ángulo diferente. El cerebro fusiona las dos imágenes, haciendo que se perciba entonces el efecto 3D. Esto permite tener una percepción extra de la información y puede ser útil para analizar las formas y el relieve de las estructuras.

El MVC dispone de un laboratorio llamado VIPL para realizar visualizaciones de datos en estéreo (Figura 4.2). Este laboratorio posee una pantalla curva de aproximadamente  $6 \times 2$  metros cubriendo un ángulo de  $135^\circ$ . La proyección se realiza mediante tres proyectores obteniendo visualizaciones en estéreo de alto rendimiento.

Los modelos 3D han sido transferidos a este laboratorio con el objetivo de ofrecer una posible aplicación futura para los analistas de imágenes basada en la estereoscopia. La visualización en estéreo de los datos puede ser interesante para realizar estudios de estructura y relieve. La Figura 4.3 muestra la persona 1 de la base de datos en el VIPL junto a su modelo 3D en tamaño real.



Figura 4.2: Vista del VIPL



Figura 4.3: Análisis en el VIPL



## Glosario de Acrónimos

<b>2D</b>	2 Dimensiones
<b>3D</b>	3 Dimensiones
<b>MVC</b>	Manchester Visualization Centre
<b>PFC</b>	Proyecto Fin de Carrera
<b>GUI</b>	Interfaz Gráfico de Usuario
<b>VRML</b>	Virtual Reality Modelling Language
<b>API</b>	Application Programming Interface
<b>POV-Ray</b>	Persistence of Vision Ray Tracer
<b>HTML</b>	Hypertext Mark-up Language
<b>CCTV</b>	Closed Circuit Televisión
<b>VIPL</b>	Visualization Immersive Projection Laboratory
<b>IT</b>	Tecnologías de la Información

## Bibliografía

- [1] Thompson T., Black S., “Forensic Human Identification”, Taylor Francis Group, LLC, 2007.
- [2] Bruce V., Young A., “In the Eye of the Beholder: The Science of Face Perception”, Oxford University Press Inc., 1998.
- [3] Proesmans M., Van Gool L., “Getting Facial Features and Gestures in 3D”, In *Face Recognition: From Theory to Applications*, Eds. H. Wechsler et al., pp. 287-309, Springer-Verlag, Berlin, 1998.
- [4] “Unit of Art in Medicine of Manchester”, The University of Manchester. <http://www.medicine.manchester.ac.uk/artinmed/>
- [5] “The University of Manchester”, The University of Manchester. <http://www.manchester.ac.uk/>
- [6] Taylor K.T., “Forensic Art and Illustration”, CRC Press LLC, 2001.
- [7] Prag J., Neave R., “Making Faces”, The Trustees of the British Museum, 1997.
- [8] Kwok Y., “Manchester Visualization Centre”, Manchester Visualization Centre, Aug 2005. <http://www.sve.man.ac.uk/mvc>
- [9] “E-frontier America, Inc. - Poser & Shade”, e-frontier America, Inc. and e-frontier, Inc, 2001-2007. <http://www.e-frontier.com/article/articleview/1156/1/281?sbss=281>
- [10] “E-frontier”, e-frontier America Inc. and e-frontier, Inc., 2001-2007. <http://www.e-frontier.com/>
- [11] “FaceShop Pro”, Software Description: DAZ Productions Inc., 2006. <http://www.daz3d.com/i.x/shop/itemdetails/-/?item=5259>. Owned by Pantomat Ltd. <sup>[12]</sup>.
- [12] “Pantomat”, Pantomat Ltd., 1994-2006. <http://www.pantomat.hu/indexeng.htm>
- [13] “Java Sun”, Sun Microsystems, Inc., 1994-2007. <http://java.sun.com/>
- [14] Flanagan D., “Java in a Nutshell”, O’Reilly, Nov 1999.
- [15] Hartman J. and Wernecke J., “The VRML 2.0 Handbook”, Silicon Graphics, Inc., 1996.
- [16] Ames A. L., Nadeau D. R. and Moreland J. L., “VRML 2.0 Sourcebook”, John Wiley & Sons, Inc, 1996.
- [17] “Cortona VRML Client”, Parallel Graphics, 2000-2007. <http://www.parallelgraphics.com/products/cortona/>
- [18] “POV-Ray - Persistence of Vision Raytracer”, Persistence of Vision Raytracer Pty. Ltd., 2003-2007. <http://www.povray.org/>
- [19] Mardia K.V., Coombes A., Kirkbridge J., Linney A., Bowie J.L., “On Statistical Problems with Face Identification from Photographs”, *Journal of Applied Statistics*, Vol.23, No.6, pp. 655-675, Journal Oxford Ltd., 1996.
- [20] Baker J., Chandran V. and Sridharan S., “Techniques for improving stereo depth maps of faces”, *International Conference on Image Processing (ICIP 2004)*, Vol.5, pp. 3279-3282, IEEE, Oct. 2004.

- [21] Gordon G.G., “3D Pose Estimation of the Face from Video”, In *Face Recognition: From Theory to Applications*, Eds. H. Wechsler et al., pp. 433-445, Springer-Verlag, Berlin, 1998.
- [22] Evans R., “Possible Art in Medicine Computer Project”, Unit of Art in Medicine of Manchester, 2006.
- [23] Hoang S., “High Performance Visualization and Virtual Reality Centre”, Manchester Visualization Centre, Oct. 2004. <http://www.sve.man.ac.uk/mvc/Service/VIPL/>